

CLAIMS

What is claimed is:

1. A method for performing a first operation within a file system in which directories and files are organized as nodes in a namespace tree, the method comprising:

associating a read-write lock with each of the nodes in the namespace tree;

acquiring a first lock on a name of one or more directories involved in the first operation;

acquiring a second lock on an entire pathname involved in the first operation;

determining whether the first lock or the second lock conflicts with third locks acquired by a second operation; and

performing the first operation when the first lock or the second lock does not conflict with the third locks, where the first, second, and third locks are read-write locks.

2. The method of claim 1, wherein the performing the first operation includes:
concurrently performing the first operation and the second operation when neither the first lock nor the second lock conflicts with the third locks.

3. The method of claim 1, wherein the first lock is a read lock.

4. The method of claim 1, wherein the second lock is one of a read lock and a write lock.

5. The method of claim 1, wherein the first operation is a read operation, the first lock is a read lock, and the second lock is a read lock.

6. The method of claim 1, wherein the first operation is a namespace modification operation, the first lock is a read lock, and the second lock is a write lock.

7. The method of claim 1, wherein the first operation is a snapshot operation, the first lock is a read lock, and the second lock is a write lock.

8. The method of claim 1, wherein the determining whether the first lock or the second lock conflicts with third locks includes:

using a lazily allocated data structure that maps pathnames to locks to determine whether the first lock or the second lock conflicts with the third locks.

9. The method of claim 1, further comprising:
serializing the first, second, and third locks when the first lock or the second lock conflicts with the third locks.

10. The method of claim 9, wherein the serializing the first, second, and third locks includes:

determining an order for the first, second, and third locks based on levels of the namespace tree involved in the first, second, and third locks and lexicographically within one of the levels of the namespace tree involved in the first, second, and third locks.

11. The method of claim 9, wherein the serializing the first, second, and third locks includes:

determining a canonical order for the first, second, and third locks, and
permitting the first and second operations to be performed based on the canonical order for the first, second, and third locks.

12. The method of claim 1, further comprising:
permitting the first and second operations to concurrently operate within a same one of the directories when neither the first lock nor the second lock conflicts with the third locks.

13. A system for performing an operation within a file system, comprising:
means for obtaining one or more first locks on one or more directory names involved in an operation;
means for obtaining a second lock on an entire pathname involved in the operation;
means for detecting whether the one or more first locks or the second lock conflict with one or more third locks acquired by another operation; and

means for executing the operation when the one or more first locks or the second lock do not conflict with the one or more third locks, the one or more first locks, the second lock, and the one or more third locks being read-write locks.

14. A file system, comprising:
a memory configured to store information regarding directories and files organized as nodes in a namespace tree; and
a processor connected to the memory and configured to:
associate a read-write lock with each of the nodes in the namespace tree,
acquire one or more first locks on names of one or more of the directories involved in a first operation,
acquire a second lock on an entire pathname involved in the first operation,
determine whether the one or more first locks or the second lock conflict with one or more third locks acquired by a second operation, and
permit the first operation to execute when the one or more first locks or the second lock do not conflict with the one or more third locks, the one or more first locks, the second lock, and the one or more third locks being read-write locks.

15. A method for performing first and second operations within a file system, comprising:
acquiring one or more first locks on one or more first directory names involved in the first operation;

acquiring one or more second locks on one or more second directory names involved in the second operation;

acquiring a third lock on a first pathname involved in the first operation;

acquiring a fourth lock on a second pathname involved in the second operation;

determining whether the first and third locks conflict with the second and fourth locks;

and

concurrently performing the first and second operations when the first and third locks do not conflict with the second and fourth locks, the one or more first locks, the one or more second locks, the third lock, and the fourth lock being read-write locks.

16. A method for concurrently performing first and second operations within a same directory, comprising:

obtaining a first lock on a sub-directory or file name within the directory by the first operation;

obtaining a second lock on a sub-directory or file name within the directory by the second operation;

determining whether the first and second locks conflict; and

concurrently performing the first and second operations when the first and second locks do not conflict, the first and second locks being read-write locks.

17. A file system, comprising:

a memory configured to store information regarding a plurality of directories and files as nodes in a namespace tree; and

a processor connected to the memory and configured to:

associate a read-write lock with each of the nodes in the namespace tree,

identify a set of the nodes involved in an operation, the identified nodes forming a pathname associated with the operation,

acquire a first one or more read-write locks, as one or more first locks, on the identified nodes,

acquire a second one of the read-write locks, as a second lock, on the pathname,

determine whether the one or more first locks or the second lock conflict with any other read-write locks, and

permit the operation to execute when the one or more first locks and the second lock do not conflict with the other read-write locks.